Asian Journal of
Science, Technology, Engineering, and Art

e-ISSN: 3025-4507
p-ISSN: 3025-5287

Index: Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/AJSTEA.v3i3.5294

Derivation of Poisson Xrama Distribution and Its Properties

Bamigbala Olateju Alao[1], Saidu Sauta Abdulkadir[2], Adesupo A. Akinrefon[3], Jibasen Danjuma[4]

[1]Federal University Wukari, Taraba States, Nigeria

[2,3,4]Modibbo Adama University, Yola, Nigeria

bamigbala@fuwukari.edu.ng

Article Info:

Submitted:	Revised:	Accepted:	Published:
Mar 1, 2025	Mar 16, 2025	Mar 28, 2025	Apr 2, 2025

Abstract

This study introduces the Poisson Xrama Distribution, a model for analyzing count data that exhibits overdispersion. By combining the Poisson distribution with the Xrama distribution, this model addresses the limitations of traditional Poisson models, which assume equidispersion. The Poisson Xrama Distribution offers enhanced flexibility in handling variance inflation, making it suitable for scenarios where standard Poisson models are insufficient. Key statistical properties, including moments, variance, skewness, kurtosis and index of dispersion measures are derived. Maximum likelihood estimation is employed for parameter estimation, providing a robust framework for practical applications. This distribution is particularly useful in fields where count data often display overdispersion, such as biology and economics, offering a promising alternative to existing distribution models.

Keywords: Overdispersion, Poisson, Equidispersion, Flexibility, Xrama

INTRODUCTION

Count data, representing the frequency of events or items, are prevalent across numerous scientific disciplines, including biology, economics, and social sciences. Such data often display dispersion characteristics: overdispersion, where the variance surpasses the mean, and underdispersion, where the variance is less than the mean. The Poisson distribution is a common tool for modeling count data due to its simplicity and its assumption of equidispersion, where the mean equals the variance. However, this assumption frequently proves inadequate, particularly in instances of overdispersion. As noted by Bektashi et al. (2022), overdispersion diminishes the effectiveness of the Poisson distribution, necessitating alternative models like the Negative Binomial (NB) distribution. The NB model introduces a dispersion parameter to account for variability beyond the Poisson model's assumptions, rendering it a popular choice for overdispersed data.

In addition to NB models, Exponential dispersion models and Conway-Maxwell-Poisson (COM-Poisson) distribution further extend flexibility by accommodating both overdispersion and underdispersion in count data (Shmueli et al., 2005; Musa & Nweze, 2021). For underdispersed data, specialized models like Generalized Poisson regression are effective. This model introduces a dispersion parameter that can handle both underdispersion and overdispersion (Famoye et al., 1997). Another robust option is the COM-Poisson model, which independently adjusts mean and variance to accommodate underdispersed scenarios (Shmueli et al., 2005). The discrete Weibull distribution model also provides adaptive capabilities for handling various dispersions relative to Poisson regression (MDPI, 2021).

Various models have been proposed to address overdispersion across different fields. For instance, the Poisson-Lindley distribution introduced by Sankaran (1970), the Poisson-Ishita distribution (PID) by Shukla and Shanker (2019), and the Poisson-Samade distribution by Aderoju et al. (2023). Additionally, Shanker and Tekie (2014) developed a quasi Poisson-Lindley distribution by compounding the Poisson distribution with a quasi Lindley distribution introduced by Shanker and Amanuel (2013). Furthermore, Nedjar and Zeghdoudi (2020) and Zeghdoudi and Nedjar (2017) introduced two compound Poisson distributions—the Poisson gamma Lindley and Poisson pseudo-Lindley distributions—by compounding the Poisson distribution with the gamma Lindley and pseudo-Lindley distributions, respectively, both proposed by Zeghdoudi and Nedjar (2016). In this research, a novel distribution tailored for overdispersed count data is introduced by compounding the Poisson distribution with the Xrama distribution. The study delves into the statistical characteristics of this new Poisson-Xrama mixture distribution.

METHODS

Poisson Xrama Distribution

The Xrama distribution is a one parameter

If $y \sim P(\theta)$, then the probability mass function (pmf) of Poisson distribution is given by

$$f(y/\theta)=\frac{\theta^{y}e^{-\theta}}{y!} \qquad y=1,2,3,\ldots \qquad \theta>0$$

Consequently, Xrama Distribution proposed by Harrison *et al.,* (2023) is given by

$$f(y;\beta)=\frac{\beta^{4}}{\left(\beta^{3}+6\right)^{2}}\left(\beta^{3}+6y^{3}+12\right)e^{-\beta y} \qquad y>0, \quad \beta>0 \tag{1}$$

If the parameter θ is assume to follow Xrama Distribution given as

$$f(\theta/\beta)=\frac{\beta^{4}}{\left(\beta^{3}+6\right)^{2}}\left(\beta^{3}+6\theta^{3}+12\right)e^{-\beta\theta} \qquad \theta>0, \quad \beta>0 \tag{2}$$

Therefore, Poisson Xrama Distribution ($PXrD(y;\beta)$) is as follows

$$P_{Y}(y)=\int_{0}^{\infty} f(y/\theta)\, f(\theta/\beta)\, d\theta$$

$$=\int_{0}^{\infty}\frac{\theta^{y}e^{-\theta}}{y!}.\frac{\beta^{4}}{\left(\beta^{3}+6\right)^{2}}\left(\beta^{3}+6\theta^{3}+12\right)e^{-\beta\theta}d\theta$$

$$P_{Y}(y)=\frac{\beta^{4}}{\left(\beta^{3}+6\right)^{2}}\left[\frac{\beta^{3}(1+\beta)^{3}+6y^{3}+36y^{2}+66y+36+12(1+\beta)^{3}}{(1+\beta)^{y+4}}\right]$$

$$y=0,1,2,\ldots;\ \beta>0 \tag{3}$$

Cumulative Density Function (CDF) of $PXrD(y;\beta)$ is obtained as:

$$F(y)=\sum_{t=0}^{y}P_t(t) \tag{4}$$

$$=\frac{\beta^4}{\left(\beta^3+6\right)^2(1+\beta)^4}\sum_{t=0}^{y}\left[\frac{\beta^3(1+\beta)^3+12(1+\beta)^3+6t^3+36t^2+66t+36}{(1+\beta)^t}\right]$$

$$F(y)=1-\left[\frac{\beta\begin{pmatrix}\beta^8+3\beta^7+3\beta^6+13\beta^5+36\beta^4+36\beta^3+156\beta^2+6\beta^2y^3\\+54\beta^2y^2+156\beta^2y+180y^2+126\beta y+216\beta+36y+144\end{pmatrix}-36}{\left(\beta^3+6\right)^2(\beta+1)^{y+4}}\right] \tag{5}$$

RESULTS

Statistical Properties of $PXrD(y;\beta)$

Moments of $PXrD(y;\beta)$

r^{th} moment of $PXrD(y;\beta)$ denoted by $\mu_{(r)}'$ is:

$$\mu_{(r)}'=E\left(y^r\right)=\sum_{y=0}^{\infty}y^rP_Y(y) \tag{6}$$

$$\mu_r=\frac{\beta^4}{\left(\beta^3+6\right)^2(1+\beta)^4}\sum_{y=0}^{\infty}y^r\left[\frac{\beta^3(1+\beta)^3+6y^3+36y^2+66y+36+12(1+\beta)^3}{(1+\beta)^y}\right] \tag{7}$$

for $r=1,2,3,4$ in equation (7), first four moments about mean of the $PXrD(y;\beta)$ are obtained below

$$\mu_1=\frac{\beta^6+12\beta^3+144}{\beta\left(\beta^3+6\right)^2}$$

$$\mu_2 = \frac{\beta^7 + 2\beta^6 + 12\beta^4 + 24\beta^3 + 144\beta + 720}{\beta^2\left(\beta^3+6\right)^2}$$

$$\mu_3 = \frac{\beta^8 + 6\beta^7 + 6\beta^6 + 12\beta^5 + 72\beta^4 + 72\beta^3 + 144\beta^2 + 2160\beta + 4320}{\beta^3\left(\beta^3+6\right)^2}$$

$$\mu_4 = \frac{\beta^9 + 14\beta^8 + 36\beta^7 + 36\beta^6 + 168\beta^5 + 432\beta^4 + 432\beta^3 + 5040\beta^2 + 25920\beta + 30240}{\beta^4\left(\beta^3+6\right)^2}$$

The Variance of $PXrD\left(y;\beta\right)$ is derived as:

$$Var\left(y\right) = \sigma^2 = \mu_2 - \left(\mu_1\right)^2 = \frac{\begin{pmatrix}\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 \\ +6048\beta^3 + 5184\beta + 5184\end{pmatrix}}{\beta^2\left(\beta^3+6\right)^4}$$

Standard deviation is given by:

$$\sigma = \frac{\sqrt{\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 + 6048\beta^3 + 5184\beta + 5184}}{\beta\left(\beta^3+6\right)^2}$$

Coefficient of Variation (CV)

$$CV = \frac{\sqrt{\sigma^2}}{\mu_1} = \frac{\sigma}{\mu_1}$$

$$CV = \frac{\sqrt{\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 + 6048\beta^3 + 5184\beta + 5184}}{\beta^6 + 12\beta^3 + 144}$$

Index of Dispersion (ID) is obtained as

$$ID = \frac{Variance}{Mean} = \frac{\sigma^2}{\mu_1}$$

$$ID = \frac{\begin{pmatrix}\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 + 6048\beta^3 \\ +5184\beta + 5184\end{pmatrix}}{\beta\left(\beta^3+6\right)^2\left(\beta^6 + 12\beta^3 + 144\right)}$$

Coefficient of Kurtosis (κ_s) is given by

$$\kappa_s = \frac{\mu_4}{(\mu_2)^2}$$

$$\kappa_s = \frac{\left(\beta^9 + 14\beta^8 + 36\beta^7 + 36\beta^6 + 168\beta^5 + 432\beta^4 + 432\beta^3 + 5040\beta^2 + 25920\beta + 30240\right)(\beta^3+6)^6}{\left(\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 + 6048\beta^3 + 5148\beta + 5184\right)^2}$$

Coefficient of Skewness (δ) is given by

$$\delta = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

$$\delta = \frac{\left(\beta^8 + 6\beta^7 + 6\beta^6 + 12\beta^5 + 72\beta^4 + 72\beta^3 + 144\beta^2 + 2160\beta + 4320\right)(\beta^3+6)^4}{\left(\beta^{13} + \beta^{12} + 24\beta^{10} + 24\beta^9 + 324\beta^7 + 648\beta^6 + 2160\beta^4 + 6048\beta^3 + 5148\beta + 5184\right)^{3/2}}$$

Generating Functions of $PXrD(y;\beta)$

The probability generating function ($P_Y(t)$), moment generating function ($M_Y(t)$) and cumulative generating function ($C_Y(t)$) of PXrD are obtained through the following.

Probability Generating Function ($P_Y(t)$)

If $Y \sim PXrD(\beta)$ then the $P_Y(t)$ is defined as:

$$P_Y(t) = \sum_{y=0}^{\infty} t^y P_Y(y)$$

$$P_Y(t) = \frac{\beta^4}{(\beta^3+6)^2(1+\beta)^4} \sum_{y=0}^{\infty} t^y \frac{\left(\beta^3(1+\beta)^3 + 6y^3 + 36y^2 + 66y + 36 + 12(1+\beta)^3\right)}{(1+\beta)^y}$$

$$P_Y(t)=\frac{\beta^4}{\left(\beta^3+6\right)^2}\left[\begin{array}{l}\frac{\left(\beta^6+3\beta^5+3\beta^4+13\beta^3+36\beta^2+36\beta+48\right)}{\left(\beta-t+1\right)^4}+\frac{\left(3\beta^5+6\beta^4-3\beta^3-36\beta^2-72\beta-36\right)t}{\left(\beta-t+1\right)^4}\\+\frac{\left(3\beta^4+3\beta^3+36\beta+36\right)t^2}{\left(\beta-t+1\right)^4}-\frac{\left(\beta^3+12\right)t^3}{\left(\beta-t+1\right)^4}\end{array}\right] \tag{8}$$

Moment Generating Function ($M_Y(t)$)

Moment generating function denoted by $M_Y(t)$ is:

$$M_Y(t)=\sum_{y=0}^{\infty}e^{ty}P_Y(y)$$

$$M_Y(t)=\frac{\beta^4}{\left(\beta^3+6\right)^2}\left[\begin{array}{l}\frac{\left(\beta^6+3\beta^5+3\beta^4+13\beta^3+36\beta^2+36\beta+48\right)}{\left(\beta-e^t+1\right)^4}+\frac{\left(3\beta^5+6\beta^4-3\beta^3-36\beta^2-72\beta-36\right)e^t}{\left(\beta-e^t+1\right)^4}\\+\frac{\left(3\beta^4+3\beta^3+36\beta+36\right)e^{2t}}{\left(\beta-e^t+1\right)^4}-\frac{\left(\beta^3+12\right)e^{3t}}{\left(\beta-e^t+1\right)^4}\end{array}\right] \tag{9}$$

Cumulative generating function denoted by $C_Y(t)$ is:

$$C_Y(t)=\phi_Y(it)$$

$$C_Y(t)=\frac{\beta^4}{\left(\beta^3+6\right)^2}\left[\begin{array}{l}\frac{\left(\beta^6+3\beta^5+3\beta^4+13\beta^3+36\beta^2+36\beta+48\right)}{\left(\beta-e^{it}+1\right)^4}+\frac{\left(3\beta^5+6\beta^4-3\beta^3-36\beta^2-72\beta-36\right)e^{it}}{\left(\beta-e^{it}+1\right)^4}\\+\frac{\left(3\beta^4+3\beta^3+36\beta+36\right)e^{2it}}{\left(\beta-e^{it}+1\right)^4}-\frac{\left(\beta^3+12\right)e^{3it}}{\left(\beta-e^{it}+1\right)^4}\end{array}\right] \tag{10}$$

Maximum Likelihood Estimation (MLE) of $PXrD(y;\beta)$

Suppose $y_1, y_2, \dots, y_n$ is a random sample of size n for $PXrD(y;\beta)$ given by

$$L(\beta / y_i) = \prod_{i=1}^{n} \frac{\beta^4}{(\beta^3+6)^2} \left[\frac{\beta^3(1+\beta)^3 + 6y^3 + 36y^2 + 66y + 36 + 12(1+\beta)^3}{(1+\beta)^{y+4}} \right]$$

log likelihood function is given by

$$\ln L(\beta / y_i) = 4n\ln\beta - 2n\ln(\beta^3+6) + \sum_{i=1}^{n} In\left(\beta^3(1+\beta)^3 + 6y_i^{\,3} + 36y_i^{\,2} + 66y_i + 36 + 12(1+\beta)^3\right) - \sum_{i=1}^{n} In\left((1+\beta)^{y_i+4}\right)$$

Differentiate w.r.t. to β, we get

$$\frac{\partial \ln(\beta / y_i)}{\partial\beta} = \frac{4n}{\beta} - \frac{6n\beta^2}{(\beta^3+6)} + \sum_{i=1}^{n} \left(\frac{3(1+\beta)^2(2\beta^3+\beta^2+12)}{\beta^3(1+\beta)^3 + 6y_i^{\,3} + 36y_i^{\,2} + 66y_i + 36 + 12(1+\beta)^3} \right) - \sum_{i=1}^{n} \left(\frac{y_i + 4n}{(1+\beta)} \right) \tag{11}$$

Substituting $\frac{\partial \ln L(\beta / y_i)}{\partial\beta} = 0$ in equation (11), This non-linear equation can be solved by any numerical iteration methods to get $\hat{\theta}_{mle}$.

CONCLUSION

This study introduces the Poisson Xrama Distribution as a novel approach to modeling overdispersed count data, addressing the limitations of traditional Poisson models. The distribution's statistical properties, including variance, standard deviation, index of dispersion, coefficient of variation, skewness, and kurtosis, are thoroughly analyzed. Maximum likelihood estimation is employed for parameter estimation. The Poisson Xrama Distribution offers enhanced flexibility in handling variance inflation, making it a valuable tool for scenarios where traditional models fall short.

REFERENCES

Aderoju, S. A., Adeniyi, I., Olaifa, J. B., & Olaosebikan, A. (2023). On Poisson-Samade Distribution: Its Applications in Modelling Count Data. *Earthline Journal of Mathematical Sciences*, *12*(2), 255-270. https://doi.org/10.34198/ejms.12223.255270

Bektashi, X., Rexhepi, S., & Limani–Bektashi, N. (2022). Dispersion of Count Data: A Case Study of Poisson Distribution and Its Limitations. *Asian Journal of Probability and Statistics*, *19*(2), 18–28. https://doi.org/10.9734/ajpas/2022/v19i230464.

Famoye, F., John T.W., & Karan P.S. (1997). On the Generalized Poisson Regression Model with an Application to Accident Data. *Journal of Data Science*, *2*(2004), 287–295.

Harrison O. Etaga, Chrisogonus K. Onyekwere, Omeje Ifunanya Lydia, Mmesoma P. Nwankwo, Dorathy O. Oramulu, Okechukwu J. Obulezi. Estimation of the Xrama Distribution Parameter Under Complete and Progressive Type-II Censored Schemes. Sch J Phys Math Stat, 2023 Dec 10(10): 203-219.

MDPI (2021). A Simple and Adaptive Dispersion Regression Model for Count Data. *Entropy*, *20*(2), 142–150.

Musa, S. I., & Nweze, N. O. (2021). Model Selection for Time Series Count Data with Over-Dispersion. *Asian Journal of Probability and Statistics*, *14*(2), 60–73. https://doi.org/10.9734/ajpas/2021/v14i230326.

Nedjar, S. & Zeghdoudi, H. (2020). New compound distribution: properties, inflated distribution, and applications, J. Agric. Stat. Sci., 16, 519-526.

Sankaran, M. (1970). The discrete Poisson-Lindley distribution. Biometrics, 26: 145- 149.

Shanker, R. & Amanuel, A. G. (2013). A new quasi Lindley distribution, International Journal of Statistics and Systems, 8, 143–156.

Shanker, R. & Tekie, A. L. (2014). A new quasi Poisson-Lindley distribution, International Journal of Statistics and Systems, 9, 87–94.

Shmueli, G., Minka, T., Kadane, J., Borle, S., & Boatwright, P. (2005). A Useful Distribution for Fitting Discrete Data: Revival of the Conway-Maxwell-Poisson Distribution. *Journal of The Royal Statistical Society: Series C*, *54*(1), 127–142.

Shukla, K.K. and Shanker, R. (2019). The Discrete Poisson-Ishita Distribution. International Journal of Statistics and Economics, 20(2):109 – 122.

Zeghdoudi, H. & Nedjar, S. (2016). A pseudo Lindley distribution and its application, Afr. Stat., 11, 923–932.

Zeghdoudi, H. & Nedjar, S. (2016). On gamma Lindley distribution: Properties and simulations, J. Comput. Appl. Math., 298, 167–174

Zeghdoudi, H. & Nedjar, S. (2017). On Poisson pseudo Lindley distribution: Properties and applications, Journal of probability and statistical science, 15, 19–28.